

VU Research Portal

Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature

van den Eerenbeemt, K.D.M.; Ostelo, R.W.J.G.; van Royen, B.J.; Peul, W.C.; van Tulder, M.W.

published in

European Spine Journal
2010

DOI (link to publisher)

[10.1007/s00586-010-1445-3](https://doi.org/10.1007/s00586-010-1445-3)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van den Eerenbeemt, K. D. M., Ostelo, R. W. J. G., van Royen, B. J., Peul, W. C., & van Tulder, M. W. (2010). Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature. *European Spine Journal*, 19(8), 1262-1280. <https://doi.org/10.1007/s00586-010-1445-3>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature

Karin D. van den Eerenbeemt · Raymond W. Ostelo ·
Barend J. van Royen · Wilco C. Peul ·
Maurits W. van Tulder

Received: 28 April 2009 / Revised: 22 April 2010 / Accepted: 9 May 2010 / Published online: 28 May 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The objective of this study is to evaluate the effectiveness and safety of total disc replacement surgery compared with spinal fusion in patients with symptomatic lumbar disc degeneration. Low back pain (LBP), a major health problem in Western countries, can be caused by a variety of pathologies, one of which is degenerative disc disease (DDD). When conservative treatment fails, surgery might be considered. For a long time, lumbar fusion has been the “gold standard” of surgical treatment for DDD. Total disc replacement (TDR) has increased in popularity as an alternative for lumbar fusion. A comprehensive systematic literature search was performed up to October

2008. Two reviewers independently checked all retrieved titles and abstracts, and relevant full text articles for inclusion. Two reviewers independently assessed the risk of bias of included studies and extracted relevant data and outcomes. Three randomized controlled trials and 16 prospective cohort studies were identified. In all three trials, the total disc replacement was compared with lumbar fusion techniques. The Charité trial (designed as a non-inferiority trial) was considered to have a low risk of bias for the 2-year follow up, but a high risk of bias for the 5-year follow up. The Charité artificial disc was non-inferior to the BAK® Interbody Fusion System on a composite outcome of “clinical success” (57.1 vs. 46.5%, for the 2-year follow up; 57.8 vs. 51.2% for the 5-year follow up). There were no statistically significant differences in mean pain and physical function scores. The Prodisc artificial disc (also designed as a non-inferiority trial) was found to be statistically significant more effective when compared with the lumbar circumferential fusion on the composite outcome of “clinical success” (53.4 vs. 40.8%), but the risk of bias of this study was high. Moreover, there were no statistically significant differences in mean pain and physical function scores. The Flexicore trial, with a high risk of bias, found no clinical relevant differences on pain and physical function when compared with circumferential spinal fusion at 2-year follow up. Because these are preliminary results, in addition to the high risk of bias, no conclusions can be drawn based on this study. In general, these results suggest that no clinical relevant differences between the total disc replacement and fusion techniques. The overall success rates in both treatment groups were small. Complications related to the surgical approach ranged from 2.1 to 18.7%, prosthesis related complications from 2.0 to 39.3%, treatment related complications from 1.9 to 62.0% and general complications from 1.0 to 14.0%.

K. D. van den Eerenbeemt
Department of Orthopaedics, EMGO Institute for Health
and Care Research, VU University Medical Center,
Amsterdam, The Netherlands

R. W. Ostelo
Department of Epidemiology and Biostatistics, EMGO Institute
for Health and Care Research, VU University Medical Center,
Amsterdam, The Netherlands

B. J. van Royen
Department of Orthopaedics, VU University Medical Center,
Amsterdam, The Netherlands

W. C. Peul
Department of Neurosurgery, Leiden University Medical Center,
Leiden, The Netherlands

M. W. van Tulder
Department of Health Sciences, EMGO Institute for Health and
Care Research, VU University, Amsterdam, The Netherlands

R. W. Ostelo (✉)
Department of Health Sciences, Faculty of Earth and Life
Sciences, VU University, De Boelelaan 1085, room U-452,
1081 HV Amsterdam, The Netherlands
e-mail: r.ostelo@vumc.nl

Reoperation at the index level was reported in 1.0 to 28.6% of the patients. In the three trials published, overall complication rates ranged from 7.3 to 29.1% in the TDR group and from 6.3 to 50.2% in the fusion group. The overall reoperation rate at index-level ranged from 3.7 to 11.4% in the TDR group and from 5.4 to 26.1% in the fusion group. In conclusion, there is low quality evidence that the Charité is non-inferior to the BAK cage at the 2-year follow up on the primary outcome measures. For the 5-year follow up, the same conclusion is supported only by very low quality evidence. For the ProDisc, there is very low quality evidence for contradictory results on the primary outcome measures when compared with anterior lumbar circumferential fusion. High quality randomized controlled trials with relevant control group and long-term follow-up is needed to evaluate the effectiveness and safety of TDR.

Keywords Degenerative disc degeneration · Low back pain · Total disc replacement · Systematic review

Introduction

Low back pain is a major health problem in Western countries [1, 2]. A variety of pathologies can cause low back pain, one of which is degenerative disc disease (DDD) [3]. It has been hypothesised that through disc dehydration, annular tears, and loss of disc height or collapse, DDD can result in abnormal motion of the segment and biomechanical instability causing pain [4–7].

When conservative treatment fails, patients and health care providers may consider other treatment options such as surgery. Although the rationale for surgery is often not clear and despite the lack of convincing evidence in the literature regarding the effectiveness of surgery in the treatment of symptomatic DDD, the number of surgical procedures performed is continually increasing [8, 9]. For a long time, lumbar fusion (arthrodesis) has been the “gold standard” surgical treatment for DDD. However, long-term results are poor and complications common [4, 10].

An alternative surgical procedure, total disc replacement, has increased in popularity. The purpose of this technique is to restore and maintain spinal segment motion, which is assumed to prevent adjacent level degeneration at the operated levels, while relieving pain [4, 11–13].

Replacing a degenerated joint instead of fusing it was considered for the spine due to the success of total knee and hip arthroplasty [5, 14, 15]. The first described total disc replacement was the Fernstorm steelball endoprosthesis in the late 1950s [16]. Since that time, multiple disc replacement prostheses have been designed for use in the lumbar spine. A large majority would never be implanted

in humans [4, 10, 17]. The first prosthesis designed to be commercially distributed as an artificial disc was initiated in 1982 by Schellnack and Buttner-Janz. Currently, many different lumbar total disc prostheses are available and approved for the European market. In the United States, American Investigational Device Exemption (IDE) trials have led to FDA approval for Charité and Prodisc prostheses.

In this article, we systematically review the available literature on the effectiveness and safety of currently available prostheses for TDR in patients with systematic DDD.

Materials and methods

Objective

The objective of this systematic review was to assess the effectiveness and safety of total disc replacement surgery in patients with chronic low back pain due to DDD. The main research questions were:

1. What is the course of DDD complaints and/or symptoms following total disc replacement surgery?
2. What is the effectiveness of total disc replacement surgery compared to other treatments?
3. What is the safety of total disc replacement surgery?

For this systematic review, we used the method guidelines for systematic reviews as recommended by the Cochrane Back Review Group [18]. Below the search strategy, selection of the studies, data extraction, risk of bias assessment, and data analysis are described in more detail. All these steps were performed by two reviewers independently (KvdE and RO) and during consensus meetings, potential disagreements between the two reviewers regarding these issues were discussed. If they were not resolved a third reviewer (MvT) was consulted.

Search strategy

An experienced librarian performed a comprehensive systematic literature search. The MEDLINE, EMBASE and COCHRANE LIBRARY databases were searched for relevant studies from 1973 to October 2008. The search strategy consisted of a combination of keywords concerning the technical procedure (e.g. disc replacement, prosthesis, implantation, discectomy, arthroplasty) and keywords regarding the anatomical features and pathology (e.g. intervertebral disc degeneration, discitis, low back pain, lumbosacral region, lumbar vertebrae). These keywords were used as MESH headings and free text words. In addition, a search was performed using the specific names

of the prostheses. The full search strategy is available upon request.

Selection of studies

The search was limited to studies published in English, German, and Dutch, because these are the languages that the review authors are able to read and to understand. Two review authors independently examined all titles and abstracts that met our search terms and reviewed full publications, when necessary. The reference section of all primary studies was inspected for additional references. For the assessment of the course of complaints and/or symptoms (research question 1), we included prospective cohort studies reporting on at least 20 cases and having a follow up period of more than 6 weeks. By definition, cohort studies do not provide information about effectiveness, so for assessment of the effectiveness (research question 2), we only included randomized controlled trials. When multiple articles were identified on the same study, but describing different follow up measurements, they were included. However, articles describing only one arm of the trial, or only describing the results of 1 centre of a multi-centre trial were excluded. For assessing the safety (research question 3), we extracted data on all reported complications from the prospective cohort studies and randomized controlled trials we included for research questions 1 and 2. Furthermore, we included overview studies on complications. Case reports were excluded.

Data extraction

Two review authors independently extracted relevant data from the included studies regarding design, population (e.g. age, gender, duration of complaints), type of total disc replacement surgery, type of control intervention (e.g. no treatment, lumbar fusion), vertebral level(s) operated on, follow-up period, and outcomes. Primary outcomes that were considered relevant were pain intensity [e.g. visual analogue scale (VAS), functional status, e.g. Roland Morris Disability Scale, Oswestry Scale (ODI), global improvement and return to work]. All ODI scores and VAS scores were converted into 0–100 scale. Other outcome measures, such as physiological outcome, radiological outcomes, and patient satisfaction were considered as secondary outcome measures.

Risk of bias (RoB) assessment

Two review authors independently assessed the risk of bias (RoB) of the included studies. Controlled trials were assessed using a criteria list recommended by the Cochrane Back Review Group [18]. The following criteria are scored

yes, no or unsure: (1) Was the method of randomization adequate? (2) Was the treatment allocation concealed? (3) Was the patient blinded to the intervention? (4) Was the care provider blinded to the intervention? (5) Was the outcome assessor blinded to the intervention? (6) Was the dropout rate described and acceptable? (7) Were all randomized participants analysed in the group to which they were allocated? (8) Are reports of the study free of suggestion of selective outcome reporting? (9) Were the groups similar at baseline regarding the most important prognostic indicators? (10) Were co-interventions avoided or similar? (11) Was compliance acceptable in all groups? (12) Was the timing of the outcome assessment similar in all groups? Criteria 11 was scored not applicable because we consider compliance not relevant for surgical interventions. If studies met at least 6 of the 12 items, the RoB was considered low. Disagreements were resolved in a consensus meeting and a third review author was consulted when necessary. Full assessment is available upon request. The overall grading of the evidence was based on the GRADE approach [19].

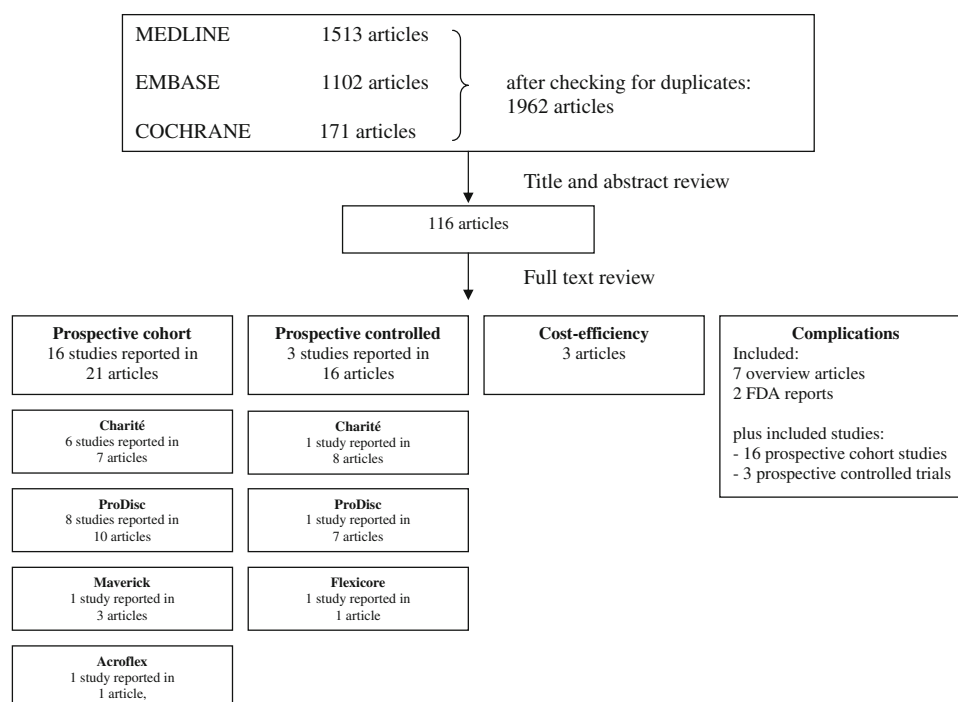
Results

Search and selection

A total of 1,962 references were identified from MEDLINE, EMBASE and the COCHRANE LIBRARY that were potentially relevant for this review on total disc replacement surgery. After checking titles and abstracts, a total of 112 full text articles were retrieved that were potentially eligible for answering all research questions. After reading full text, 21 articles reporting on 16 studies were relevant for answering research question 1, and 16 articles reporting on 3 studies were relevant for research question 2. Seven overview articles for answering research question 3. Figure 1 shows the search strategy process in a flow diagram. Reviewing the reference lists of these articles resulted in no additional studies.

Type of studies

For assessing the course of DDD complaints and/or symptoms (research question 1) 16 prospective cohort studies were included, describing four different devices, 6 for Charité [20–26], 8 for Prodisc [27–36], 1 for Maverick [37–39], and 1 for Acroflex [40]. For assessing the effectiveness of total disc replacement (research question 2), we identified three randomized controlled trials, all conducted in the USA in order to get FDA approval. Each trial examined a different prostheses for TDR (Charité, and ProDisc and Flexicore) and all used fusion (although

Fig. 1 Flow diagram

different types) as control intervention. No studies comparing TDR surgery to other treatments were found. Two trials (Charité and ProDisc trial) were described in multiple articles [6, 10, 13, 41–53]. Given our inclusion and exclusion criteria, we finally included two articles describing the Charité trial; one reporting the 24 months follow up [10] and the other the 5-year follow up [51]. However, 5-year follow-up results were only available of 57% of the originally randomized population. We included one article for the ProDisc trial reporting the 24-month follow-up results [52]. The Flexicore trial was described in one article, which should be considered as preliminary results as the final results of this trial have not yet been published [53] (Table 1).

Assessment of risk of bias

After assessing, the risk of bias of the controlled trials there was 20% disagreement between the two review authors. Consultation of the third reviewer was not necessary because disagreements were resolved in a consensus meeting. For the 24 months follow up, the reporting on the Charité trial was considered to have a low risk of bias. However, the reporting on the 5-year follow up was considered to have a high risk of bias. The reporting on the ProDisc trial was considered to have a high risk of bias as it only met 4 out of the 11 risk of bias criteria. The Flexicore study was also considered to have a high risk of bias as it only met 2 out of 11 risk of bias criteria (Table 2). By

design, the prospective cohort studies were not only included in the effectiveness analysis, but also used to describe the course of DDD complaints and/or symptoms after undergoing a TDR surgery.

Outcomes

(1) *What is the course of DDD complaints and/or symptoms following total disc replacement surgery?*

Charité® (Table 3) The Charité prostheses is the first total disc prostheses, developed by Butter-Janzen and Schellnack at the Charité clinic in former East Germany. The CharitéIII became commercially available for the first time in the late 1980s [54, 55]. Lemaire et al. reported 2 articles, respectively, with 51 months follow-up in 1997 [20] and with 11.3 years follow up in 2005 [23] on the same population. These articles report a good or excellent clinical result, respectively, in 85 and 90%. Several other prospective cohort studies report positive results as well on VAS improvement (range 16–66 points), ODI improvement (range 14–51%) and patients' satisfaction (range 69–92%) [21–26].

ProDisc® (Table 3) The ProDiscI was developed in France by Marnay, who operated on 64 patients and performed a single or multi-level total disc replacement in the beginning of the 1990s [29, 54, 55]. Fifty-five patients were available for follow-up after average 8.7 years. 82.6% of

Table 1 Prospective controlled studies

Study/author	Inclusion/exclusion criteria	Interventions/demografie	Follow up: duration and outcome
Sasso et al. [53] Randomized N = 76	<p>Inclusion criteria: single-level symptomatic DDD at L1–S1 confirmed by radiology, more axial than radicular pain, VAS ≥ 40 and ODI ≥ 40, failed ≥ 6 mo of conservative treatment</p> <p>Exclusion criteria: previous bilateral or unilateral lumbar decompression, microdiscectomy, lumbar fusion, spondylolysis, isthmic spondylolisthesis, moderate to severe spinal stenosis, lumbar scoliosis, facet joint arthritic changes, significant motion segment instability, Paget disease, osteopenia (including osteoporosis or osteomalacia), metabolic bone disease</p>	<p>Index: TDR Type: FlexiCore® Level: L5–S1 ($n = 32$) L4–L5 ($n = 12$) N = 44 ♀21 ♂23 (50 randomized) mean 36 years</p> <p>Control: anterior lumbar circumferential fusion Type: femoral ring allograft, autogenous iliac crest bone, pedicle screw Level: L5–S1 ($n = 17$) L4–L5 ($n = 5$) L4–L5–S1 ($n = 1$) N = 23 ♀13 ♂10 (26 randomized) mean 41 years</p>	<p>Follow up I: 24 months (6 weeks, 3, 6, 12, 24 months), 75% lost to follow up C: 24 months (6 weeks, 3, 6, 12, 24 months), 70–35 lost to follow up Pain (VAS) I: pre-op.: 86, follow up: 16, difference: –70 C: pre-op.: 82, follow up: 20, difference: –62 Functional status (ODIv2.0) I: pre-op: 62%, follow up: 6%, difference –56% C: pre-op: 58%, follow up: 12%, difference –46% Complications I: $n = 10$ (22.7%), C: $n = 10$ (43.5%) Re-operations I: $n = 5$ (11.4%), C: $n = 6$ (26.1%)</p>
Zigler et al. [52] Randomized N = 236	<p>Inclusion criteria: single-level symptomatic DDD at L3–S1 confirmed by radiology, back and/or leg (radicular) pain, ODI ≥ 40 (20/50), failed ≥ 6 mo of conservative treatment</p> <p>Exclusion criteria: prior fusion surgery, current or prior fracture at L3–S1, osteoporosis, spondylolisthesis, facet joint disease or degeneration, spinal stenosis, back or leg pain of unknown etiology, Paget's disease, osteomalacia, any other metabolic bone disease, allergy to device</p>	<p>Index: TDR Type: ProDisc®-L Level: L5–S1 ($n = 104$) L4–L5 ($n = 54$) L3–L4 ($n = 3$) N = 161 ♀79 ♂82 mean 38.7 years, SD 8.0</p> <p>Control: anterior lumbar circumferential fusion Type: femoral ring allograft, autogenous iliac crest bone, pedicle screw Level: L5–S1 ($n = 50$) L4–L5 ($n = 22$) L3–L4 ($n = 3$) N = 75 ♀41 ♂34 mean 40.4 years, SD 7.6</p>	<p>Follow up I: 24 months (6 weeks, 3, 6, 12, 18, 24 months), 9.0% lost to follow up C: 24 months (6 weeks, 3, 6, 12, 18, 24 months), 11.5% lost to follow up Pain (VAS) $P = 0.08$ I: pre-op: 76, follow up: 37 (SD30.1), difference: –39 C: pre-op: 75, follow up: 43 (SD31.6), difference: –32 Functional status (ODI, chiropractor revised version), $P = 0.0551$ I: pre-op: 63.4% (SD12.6), follow up: 34.5% (SD24.8), difference –28.9% C: pre-op: 62.7% (SD10.3), follow up: 39.8% (SD24.3), difference –22.9% Clinical success (ODI improvement $\geq 25\%$) $P = 0.0396$ I: 69.1%, C: 54.9% Functional status (SF-36 improvement) $P = 0.094$ I: 79.2%, C: 70.0% Return to work $P = 0.0485$ I: pre-op: 83.5%, follow up: 92.4%, C: pre-op: 78.1%, follow up: 85.1% Patient Satisfaction (VAS) $P = 0.015$ I: 76.7 (SD 29.2), C: 67.3 (SD31.5) Surgery again? $P = 0.1304$ I: 81%, C: 69% Overall success, sponsor definition (10 endpoints*) $P = 0.0053$ I: 63.5%, C: 45.1% Overall success, FDA definition (10 endpoints**) $P = 0.0438$ I: 53.4%, C: 40.8% Complications I: $n = 10$ (6.2%), C: $n = 5$ (6.7%) Re-operations I: $n = 6$ (3.7%), C: $n = 4$ (5.4%)</p>

Table 1 continued

Study/author	Inclusion/exclusion criteria	Interventions/demografie	Follow up: duration and outcome
Blumenthal et al. [10] <i>Randomized</i> <i>N</i> = 304	<i>Inclusion criteria:</i> single level symptomatic DDD at L4–S1 confirmed by discography, back and/or leg pain without radiculopathy, VAS ≥ 40 and ODI ≥ 30 , failed ≥ 6 mo of conservative treatment <i>Exclusion criteria:</i> Previous thoracic or lumbar fusion, current or prior, fracture at L4–S1, osteoporosis, spondylolisthesis > 3 mm, scoliotic deformity $> 11^\circ$	Index: TDR Type: Charité® Level: L4–L5 (<i>n</i> = 61) L5–S1 (<i>n</i> = 144) <i>N</i> = 205 ♀92 ♂113 mean 39.6 years (8.16), range 19–60 Control: anterior lumbar interbody fusion Type: BAK Level: L4–L5 (<i>n</i> = 32) L5–S1 (<i>n</i> = 67) <i>N</i> = 99 ♀55 ♂44 mean 39.6 years (9.07), range 20–60	Follow up I: 24 months, 8.5% lost to follow up C: 24 months 10.8% lost to follow up Pain (VAS) <i>P</i> = 0.1074 I: pre-op: 72.0, follow up: 31.2, difference: –40.6 C: pre-op: 71.8, follow up: 37.5, difference: –34.1 Functional status (ODIv1.0) <i>P</i> = 0.267 I: pre-op: 50.6%, follow up: 26.3%, difference –24.3%, C: pre-op: 52.1%, follow up: 30.5%, difference –21.6% Clinical success (ODI improvement $\geq 25\%$) <i>P</i> = 0.0038 I: 63.9%, C: 50.5% Return to work <i>P</i> = 0.6329 I: pre-op: 53.2%, follow up: 62.4%, difference 9.2%, C: pre-op: 57.6%, follow up: 65%, difference 7.4% Patient Satisfaction <i>P</i> = 0.0011 I: 73.7%, C: 53.1% Who would choose same treatment? <i>P</i> = 0.0062 I: 69.9%, C: 50.0% Overall success (4 criteria***) <i>P</i> < 0.0001 ^a I: 57.1%, C: 46.5% Complications I: <i>n</i> = 56 (27.3%), C: <i>n</i> = 46 (46.5%) Re-operations I: <i>n</i> = 11 (5.4%), C: <i>n</i> = 9 (9.1%) Follow up I and C: 5 years, 43% lost to follow up Pain (VAS) <i>P</i> = NS I: pre-op: 72.0, follow up: 30, difference: –42 C: pre-op: 71.8, follow up: 30, difference: –41.8 Clinical success (ODI improvement ≥ 15 point) <i>P</i> = 0.8443 I: follow up: 68%, C: follow up: 65% Return to work <i>P</i> = 0.0403 I: 65.6%, C: 46.5% Overall success (4 criteria***) <i>P</i> = 0.0359 ^a I: 57.8%, C: 51.2% Complications not mentioned Re-operations (2 to 5 years) I: <i>n</i> = 2 (2.2%), C: <i>n</i> = 0
Guyet et al. [51] <i>Randomized</i> <i>N</i> = 233	<i>Inclusion criteria:</i> single level symptomatic DDD at L4–S1 confirmed by discography, back and/or leg pain without radiculopathy, VAS ≥ 40 and ODI ≥ 30 , failed ≥ 6 mo of conservative treatment <i>Exclusion criteria:</i> Previous thoracic or lumbar fusion, current or prior, fracture at L4–S1, osteoporosis, spondylolisthesis > 3 mm, scoliotic deformity $> 11^\circ$	Index: TDR Type: Charité® Level: L4–L5 (<i>n</i> = 26) L5–S1 (<i>n</i> = 64) <i>N</i> = 90 ♀43 ♂47 mean 40.0 years (8.58), range 19–60 Control: anterior lumbar interbody fusion Type: BAK Level: L4–L5 (<i>n</i> = 10) L5–S1 (<i>n</i> = 33) <i>N</i> = 43 ♀19 ♂24 mean 38.8 years (8.69), range 25–55	

* 10 points for clinical success, sponsor definition: $\geq 15\%$ improvement in ODI, device success, neurologic success, SF-36 improvement, six radiographic outcomes: no migration, no subsidence, no radiolucency, no loss of disc height, fusion status and ROM

** 10 points for clinical success, FDA definition: ≥ 15 -point improvement in ODI, device success, neurologic success, SF-36 improvement, six radiographic outcomes: no migration, no subsidence, no radiolucency, no loss of disc height, fusion status and ROM

*** 4 criteria: $\geq 25\%$ improvement in ODI, device failure, major complications, neurological deterioration

^a Blackwelder's test of noninferiority

VAS (scale 0–100), ODI (scale 0–100), NS not specified

Table 2 Methodological quality prospective controlled studies

Trial		A1	B2	C3	C4	C5	D6	D7	E8	F9	F10	F11 ^a	F12	Risk of bias
FlexiCore [®]	Sasso et al. [53]	U	U	N	N	N	N	N	U	Y	U	NA	Y	2/11 (high)
ProDisc [®]	Zigler et al. [52]	Y	U	N	N	N	N	N	Y	Y	U	NA	Y	4/11 (high)
Charité [®]	Blumenthal et al. [10]	Y	Y	N	N	N	Y	Y	Y	Y	U	NA	Y	7/11 (low) ^b
	Guyer et al. [51]	Y	Y	N	N	N	Y	N	Y	Y	U	NA	Y	6/11 (low) ^c
		Y	Y	N	N	N	N	U	Y	Y	U	NA	Y	5/11 (high) ^d

U unsure, Y yes, N no, NA not applicable

^a Criteria 11 was scored

^b For overall measures, 2 year follow-up

^c For VAS en ODI outcome, 2 year follow-up

^d For 5 year follow-up

the patients were “completely satisfied” or “satisfied” with the results [29]. ProDiscII, the second generation ProDisc, is reported on in several publications with follow up range 3 months to 2 years [27, 28, 30–36]. Primary outcome results suggest being positive, VAS improvement (range 40–62 points) and ODI improvement (range 21–48%). Moreover, the majority of patients seem to be satisfied with the results (range 79–100%).

Maverick[®] (Table 3) Huec et al. published several studies on the Maverick device [37–39]. At 2-year follow-up improvement was reported in VAS for low back pain and leg pain, decreasing by 44 and 18 points, respectively. Functional status improved as ODI score decreased 20.7%, and an overall improvement in functional status of 25% occurred in 75% of the patients. Since 2003, a prospective controlled trial has been ongoing in the USA [54].

Acroflex[®] (Table 3) Fraser et al. [40] conducted two pilot studies and combined the results. The endplates were changed for the second study because of device failure. For the whole group, the functional impairment (ODI) improved from 14.9% and the Low Back Pain Score (LBPS) improved from 17.7 to 33 at 24 months follow-up. 50% of the patients were not working because of their back condition. Due to detection of mechanical failure, the randomized controlled trial has not been carried out.

In conclusion, many studies suggest pain relief, improvement in functional status and patient satisfaction after TDR surgery. The overall outcome is positive, reduction of pain intensity (range 16–66 points) and improvement of functional impairment (range 14–51%). Moreover, the majority of patients seem to be satisfied with the results (range 69–100%). Unfortunately, detailed information on how outcomes were measured was often lacking. Although outcome results from observational studies suggest a positive course after TDR surgery, a

drawback is that a significant amount of complications was reported as well (which will be described later), and a control group was lacking in these studies.

(2) *What is the effectiveness of total disc replacement surgery compared to other treatments?*

Charité[®] trial (Table 1) The Charité trial [10, 51], which was designed as a non-inferiority trial, randomized 304 patients to either TDR with the Charité III disc ($n = 205$) or anterior interbody fusion with BAK cage ($n = 99$) with a follow-up of 2 and 5 years. The primary outcomes were pain (VAS), functional impairment (ODI), overall clinical success (defined by using four criteria: $\geq 25\%$ improvement in ODI, device failure, major complications, and neurological deterioration). As a secondary outcome, patient satisfaction was measured. The improvements on pain intensity (-40.6 vs. -34.1) and functional impairment (24.3 vs. 21.6%), for the TDR and the BAK, respectively, did not differ significantly at 2-year follow up. The overall clinical success (indeed statistically tested on non-inferiority) revealed that the Charité group was non-inferior to the lumbar fusion group (57.1 vs. 46.5% ; $P < 0.0001$. P value based on the Blackwelder’s test for equivalence). Patient satisfaction was significantly better in the Charité group (73.7%) compared to the control group (53.1%) ($P < 0.002$). 5-year results, based on only 57% of the randomized patients and with a high risk of bias, were broadly in line with the 2-year results. At 5-year follow-up, outcomes on the composite score of clinical success showed that the Charité was non-inferior to the lumbar fusion group (57.8 vs. 51.2% ; $P < 0.04$. P value based on the Blackwelder’s test for equivalence) [51]. There were no statistically significant differences in functional impairment and pain intensity. In conclusion, there is low quality evidence (based on one study only with a low risk of bias) that there are no clinically relevant differences on the

Table 3 Prospective cohort studies

Study/author	Indication in/exclusion criteria	Demografy/ type of disc for TDR/ Level of intervention	Follow up: duration and outcome
Charité®			
Zeegers et al. [21]	<p><i>Indication/diagnosis:</i></p> <p>discopathy (40)</p> <p>post-discectomy discopathies (29)</p> <p>discopathy with possible disc protrusion (6)</p> <p><i>Exclusion criteria:</i></p> <p>predominant symptoms or deficits in the legs related to involvement of the nerve roots, spondylolisthesis, spinal stenosis, altered posterior elements, infection, metabolic bone diseases (e.g. osteoporosis, osteomalacia), severe scarring after previous surgery, insufficient motivation of the patient.</p>	<p>$N = 50 \text{ ♀} 30 \text{ ♂} 20$</p> <p>mean 43 years (range 24–59 years)</p> <p>Type: Charité III®</p> <p>Level:</p> <p>1-level L5/S1 ($n = 8$), L4/L5 ($n = 16$), L3/L4 ($n = 5$)</p> <p>2-level L4/L5/S1 ($n = 12$)</p> <p>L3/L4/L5 ($n = 5$)</p> <p>L5/S1 and L3/L4 ($n = 1$)</p> <p>3-level L3/L4/L5/S1 ($n = 3$)</p>	<p>Follow up: 24 months</p> <p>8–32% lost to follow up</p> <p>Low Back pain (improvement): 65% (30/46)</p> <p>Leg pain (improvement): 64% (27/42)</p> <p>Return to work: 81% (35/43)</p> <p>Narcotic use (decrease): 30% (15/34)</p> <p>Patient Satisfactory (no regret of surgery): 83% (38/46)</p> <p>Overall clinical results (Stauffer Coventry Scale): Good and fair: 70% (32/46), Poor: 30% (14/46)</p> <p>Complication: $n = 52$ (in 30 patients = 60%)</p> <p>Re-operation: ($n = 7$) 14%</p>
Regan [22]	<p><i>Indication/Diagnosis:</i></p> <p>DDD, symptomatic</p> <p><i>Inclusion criteria</i></p> <p>age between 18–60 years, single level DDD, contained herniated nucleus pulposus, paucity of facet joint degeneration, primary complaint of back pain, failed conservative treatment for at least 6 months, a minimum disc height of 4 mm, scarring/thickening of the annulus fibrosis with osteophytes indicating osteoarthritis.</p> <p><i>Exclusion criteria:</i></p> <p>previous thoracic or lumbar fusion, previous surgery on the effected level (except prior discectomy, nondestabilizing laminectomy/otomy without facetectomy, or nucleosus), osteoporosis or metabolic bone disease, radicular leg pain, lumbar scoliosis, stenosis, segmental instability, spondylosis, spondylolisthesis, spinal neoplasm, active systemic or local infection, facet joint arthrosis, pregnancy, history of chronic steroid use, arachnoiditis, metal allergy, autoimmune disorders, psychosocial disorder, previous retroperitoneal or > 3 intra-abdominal operations</p>	<p>$N = 100 \text{ ♀} 46 \text{ ♂} 54$</p> <p>mean 43 years (range 24–59 years)</p> <p>Type: Charité III®</p> <p>Level:</p> <p>1-level L5/S1 ($n = 68$)</p> <p>L4/L5 ($n = 32$)</p>	<p>Follow up: 6 months (6 weeks, 3, 6 months)</p> <p>0% lost to follow up</p> <p>Pain (VAS): pre-op: 73.2 (SD14.5), follow up: 39.2 (SD26.4), difference: –34.0</p> <p>Functional status (ODI): pre-op 53.4% (SD13.4) follow up: 37.6% (SD18.6), difference: –15.8%</p> <p>Complication: $n = 9$ (10%)</p> <p>Re-operation: $n = 7$ (7.0%)</p>
Lemaire et al. [20, 23]	<p><i>Indication/Diagnosis:</i></p> <p>DDD, symptomatic</p> <p><i>Inclusion criteria</i></p> <p>failed conservative treatment</p> <p><i>Exclusion criteria:</i></p> <p>obesity, prior fusion, instability such as spondylolisthesis, deformity, radicular pain symptomology, presence of facet arthrosis</p>	<p>$N = 100 \text{ ♀} 59 \text{ ♂} 41$</p> <p>mean 40 years (range 24–51 years)</p> <p>Type: Charité III®</p> <p>Level:</p> <p>1-level: ($n = 54$)</p> <p>2-level: ($n = 45$)</p> <p>3-level: ($n = 1$)</p> <p>L5/S1 ($n = 72$)</p> <p>L4/L5 ($n = 69$)</p> <p>L3/L4 ($n = 6$)</p>	<p>Follow up: mean 11.3 months (range 10–13.4 years)</p> <p>7% lost to follow up ($n = 107$ originally included)</p> <p>Return to work: 91.6% (87/95)</p> <p>Overall clinical results (Stauffer Coventry Scale): Excellent: 62%, Good: 28%, Poor: 10%</p> <p>Complication: $n = 16$ (16%)</p> <p>Re-operation: $n = 5$ (5%)</p>
Ross et al. [24]	<p><i>Indication/Diagnosis:</i></p> <p>DDD, symptomatic</p> <p><i>Inclusion criteria</i></p> <p>failed conservative treatment, ODI > 30%</p> <p><i>Exclusion criteria:</i></p> <p>spondylolisthesis</p>	<p>$N = 160 \text{ ♀} 98 \text{ ♂} 62$</p> <p>mean 46 years (range 27–73 years)</p> <p>Type: Charité III®</p> <p>Level:</p> <p>L5/S1 ($n = 114$)</p> <p>L4/L5 ($n = 92$)</p> <p>L3/L4 ($n = 20$)</p>	<p>Follow up: mean 79 months (range 31–161 months)</p> <p>23.1% lost to follow up</p> <p>Pain (VAS): pre-op*: 63, follow up**: 47, difference: –16</p> <p>Functional status (ODI): pre-op*: 51% (26–90) follow up**: 37% (0–90) difference: –14%</p> <p>Patient Satisfactory**: “much better” 41%, “better” 28%, “same as before” 11%, “worse than before” 20%</p> <p>Complication: $n = 53$ (33.1%)</p> <p>Re-operation: $n = 12$ (7.5%)</p> <p>*($n = 77/160$), **($n = 123/160$)</p>

Table 3 continued

Study/author	Indication/in/exclusion criteria	Demografy/type of disc for TDR/Level of intervention	Follow up: duration and outcome
Gioia et al. [25]	<i>Indication/Diagnosis:</i> DDD, symptomatic <i>Exclusion criteria:</i> disc herniation, spinal canal stenosis, spondylolysis, osteoporosis.	$N = 36 \text{ ♀} 23 \text{ ♂} 13$ mean 40 years (range 32–49 years) Type: Charité III® Level: 1-level: ($n = 28$) 2-level: ($n = 7$) 3-level: ($n = 1$) L5/S1 ($n = 15$) L4/L5 ($n = 11$) L3/L4 ($n = 2$)	Follow up: mean 6.9 years (range 5–9 years) lost to follow up not mentioned Pain (VAS): pre-op: 80 (75–83), follow up: 14 (9–19), difference: –66 Functional status (ODI): pre-op: 44% (39.3–48.7), follow up: 9% (5.4–12.2), difference: –35% Patient Satisfactory: “excellent” 72%, “good” 20%, “inadequate” 4%, “poor” 4% Surgery again?: ‘yes’ 92% (33/36) Complication: $n = 11$ (30.6%) Re-operation: $n = 5$ (13.9%)
Warachit [26]	<i>Indication/Diagnosis:</i> DDD, symptomatic	$N = 43 \text{ ♀} 17 \text{ ♂} 26$ mean 42 years (range 23–54 years) Type: Charité III® Level: 1-level: ($n = 36$) 2-level: ($n = 7$) L5/S1 ($n = 16$) L4/L5 ($n = 33$) L3/L4 ($n = 1$)	Follow up: 3 years lost to follow up not mentioned Pain (VAS): pre-op: 74, follow up: 13, difference: –61 Functional status (ODI): pre-op: 60.9%, follow up: 9.8% difference: –51.1% Complication: $n = 3$ (7.0%) Re-operation: $n = 1$ (2.3%)
Prodisc® Bertagnoli et al. [27]	<i>Indication/Diagnosis:</i> Degenerative disc disease (DDD) ($n = 67$) Failed disc surgery syndrome ($n = 35$) Transition zone syndrome (TZS) ($n = 6$) <i>Exclusion criteria:</i> Severe osteoporosis, physiological dysfunction, history of previous infection, severe posterior element pathologies, fracture of the vertebra, tumor.	$N = 108 \text{ ♀} 50 \text{ ♂} 58$ mean 42 years (range 34–65 years) Type: ProDisc II® Level: 1-level L5/S1 ($n = 61$) L5/L6 ($n = 3$) L4/L5 ($n = 31$) L3/L4 ($n = 7$) L2/L3 ($n = 3$) 2-level L4/L5/S1 ($n = 10$) L2/L3 and L4/L5 ($n = 1$) 3-level L3/L4/L5/S1 ($n = 2$)	Follow up: range 3 months–2 years lost of follow up not mentioned exact, at least 50% lost to follow up at 1 year Return to work: same level 64.8% ($n = 35$), lower level 31.5% ($n = 17$), unable 3.7% ($n = 2$) Resume daily activity: average 2.3 wk (range 1.5–3.2 wk) Overall: excellent 90.8% ($n = 98$), good 7.4% ($n = 8$), fair 1.8% ($n = 2$), poor 0% ($n = 0$) Complication: $n = 1$ (1.0%) Re-operations: $n = 0$
Tropiano et al. [28]	<i>Indication/Diagnosis:</i> DDD, symptomatic ($n = 33$) Failed spine surgery ($n = 20$) <i>Inclusion criteria</i> Failed ≥ 6 months of conservative treatment <i>Exclusion criteria</i> Associated facet degeneration, history of abdominal or retroperitoneal surgery near planned anterior approach, osteoporosis, osteopenia, structural spinal deformities, absence (postoperative) of posterior elements, chronic disease of major organ system, history of local infection, pregnancy	$N = 53 \text{ ♀} 35 \text{ ♂} 18$ mean 45 years (range 28–68 years) Type: ProDisc II® Level: 1-level L5/S1 ($n = 27$), L4/L5 ($n = 13$), 2-level L4/L5/S1 ($n = 9$) L3/L4/L5 ($n = 2$) 3-level L3/L4/L5/S1 ($n = 2$)	Follow up: mean 1.4 years (range 1–2 years) lost to follow up not mentioned Pain lumbar (VAS): pre-op: 74 (SD 25), follow up: 13 (SD 18), difference: –61 Pain radicular (VAS): pre-op: 67 (SD 30), follow up: 19 (SD 26), difference: –48 Functional status (ODI): pre-op: 56% (SD 8.2), follow up: 14% (SD 7.4), difference: –42% Return to work: “normal”: 71.7%, “slightly limited”: 28.3%, “impossible”: 13.2% Patient Satisfaction: “entirely satisfied”: 86.8%, “satisfied”: 13.2%, “not satisfied”: 0% Complication: $n = 5$ (9.4%) Re-operations: $n = 3$ (5.7%)

Table 3 continued

Study/author	Indication/in/exclusion criteria	Demography/type of disc for TDR/Level of intervention	Follow up: duration and outcome
Tropiano et al. [29]	<p><i>Indication/Diagnosis:</i> DDD, symptomatic</p> <p><i>Inclusion criteria:</i> Single- and multi-level symptomatic DDD confirmed by radiology, severe lumbar pain, failed ≥ 6 months of conservative treatment</p> <p><i>Exclusion criteria:</i> Facet arthrosis, central or lateral recess stenosis, osteoporosis, sagittal or coronal plane deformity, postoperative absence of posterior elements, sequestered herniated nucleus pulposis</p>	<p>$N = 55$ ♀25 ♂30 mean 46 years (range 25–65 years)</p> <p>Type: ProDisc I®</p> <p>Level: 1-level L5/S1 ($n = 10$), L4/L5 ($n = 17$), L4/L5* ($n = 6$), L3/L4 ($n = 2$) 2-level L3/L4/L5 ($n = 1$) L3/L4/L5* ($n = 1$) L4/L5/S1 ($n = 15$) 3-level L3/L4/L5/S1 ($n = 3$) *with concomitant L5–S1 arthrodesis</p>	<p>Follow up: mean 104 months (range 84–128 months)</p> <p>14.1% lost to follow up ($n = 64$ originally included)</p> <p>Low-back pain (severe or moderate) pre-op: 94.5%, follow-up: 43.6%, difference: –50.9%</p> <p>Lower-limb pain (severe or moderate) pre-op: 85.5%, follow-up: 21.8%, difference: –63.7%</p> <p>Impairment (substantially or limited or impossible) pre-op: 83.6%, follow-up: 20.0%, difference: –63.6%</p> <p>Functional status, Modified Stauffer-Coventry score (scale 0–20) pre-op: 7.0 (SD 3.3), follow up: 16.1 (SD 2.8), difference: ± 9.1</p> <p>Patient Satisfaction: “completely satisfied” 63.6%, “satisfied” 27.3%, “not satisfied” 9.0%</p> <p>Overall: excellent 60%, good 14.5%, poor 25.5%</p> <p>Complications: $n = 10$ (18.2%) Re-operations: $n = 3$ (5.5%)</p>
Bertagnoli et al. [30]	<p><i>Indication/Diagnosis:</i> Single-level DDD, symptomatic</p> <p><i>Inclusion criteria:</i> Single-level symptomatic DDD at L3–S1 confirmed by radiology, back and/or leg (radicular) pain, failed ≥ 9 months of conservative treatment, age between 18 and 60 years</p> <p><i>Exclusion criteria:</i> patients with spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, pregnancy, facet arthrosis, inadequate vertebral endplate size, more than one level of spondylosis, neuromuscular disease, Worker’s Compensation, spinal litigation, body mass index greater than 35, and/or any isthmic or degenerative spondylolisthesis greater than Grade 1</p>	<p>$N = 104$ ♀57 ♂47 median 48 years (range 36–60 years)</p> <p>Type: ProDisc II®</p> <p>Level: 1-level L5/S1 ($n = 80$), L4/L5 ($n = 17$) L3/L4 ($n = 7$)</p>	<p>Follow up: 24 months (6 weeks, 3, 6, 12, 24 months), 12% lost to follow up ($n = 118$ originally included)</p> <p>Pain (VAS): pre-op: 75, follow up: 30, difference: –45</p> <p>Back Pain: No pain: pre-op: 0%, follow up: 32.0%, difference: +32.0%</p> <p>Occasional pain: pre-op: 15.3%, follow up: 59.2%, difference: +43.9%</p> <p>Regular pain: pre-op: 84.6%, follow up: 9.0%, difference: –75.6%</p> <p>Radicular Pain: No pain: pre-op: 11.9%, follow up: 62.6%, difference: +50.7%</p> <p>Occasional pain: pre-op: 45.5%, follow up: 29.5%, difference: –16.0%</p> <p>Regular pain: pre-op: 42.6%, follow up: 8.8%, difference: –33.8%</p> <p>Functional status (ODI): pre-op: 54% follow up: 29%, difference: –25%</p> <p>Return to work: Full-time: threefold increase, Part-time: fourfold increase, not working: fivefold decrease</p> <p>Patient Satisfactory: “completely satisfied” 58.3%, “satisfied” 38.8%, “unsatisfied” 3.1%</p> <p>Complication: $n = 5$ (4.2%) Re-operations: $n = 1$ (1.0%)</p>

Table 3 continued

Study/author	Indication/in/exclusion criteria	Demografy/type of disc for TDR/Level of intervention	Follow up: duration and outcome
Bertagnoli et al. [31]	<p><i>Indication/Diagnosis:</i> Multi-level DDD, symptomatic</p> <p><i>Inclusion criteria:</i> Multi-level symptomatic DDD at L1–S1 confirmed by radiology, back and/or leg (radicular) pain, failed ≥ 9 months of conservative treatment, age between 18 and 60 years</p> <p><i>Exclusion criteria:</i> Patients with spinal stenosis, osteoporosis, prior fusion surgery, chronic infections, metal allergies, pregnancy, facet arthrosis, inadequate vertebral endplate size, more than one level of spondylosis, neuromuscular disease, Worker's Compensation, spinal litigation, body mass index greater than 35, and/or any isthmic or degenerative spondylolisthesis greater than Grade 1</p>	<p>$N = 25$ ♀10 ♂15 median 51 years (range 30–60 years) Type: ProDisc II® Level: 2-level L2/L3 and L4/L5 ($n = 1$) L3/L4 and L5/S1 ($n = 1$) L3/L4/L5 ($n = 5$) L4/L5/S1 ($n = 8$) 3-level L3/L4/L5/S1 ($n = 10$)</p>	<p>Follow up: 24 months (6 weeks, 3, 6, 12, 24 months), 13.8% lost to follow up ($n = 29$ originally included) Pain (VAS): pre-op: 83 (60–100), follow up: 21 (0–60), difference: –62 Back Pain: No pain: pre-op: 0%, follow up: 56%, difference: +56% Episodic pain: pre-op: 8%, follow up: 36%, difference: +28% Regular pain: pre-op: 92%, follow up: 8%, difference: –84% Radiculair Pain: No pain: pre-op: 20%, follow up: 67%, difference: +47% Episodic pain: pre-op: 28%, follow up: 33%, difference: +5% Regular pain: pre-op: 52%, follow up: 0%, difference: –52% Functional status (ODI): pre-op: 65.0% (42–92) follow up: 21.6% (0–48), difference: –43.4% Return to work: Full-time: fivefold increase, Part-time: twofold increase, not working: fourfold decrease Patient Satisfactory: “completely satisfied”: 75%, “satisfied”: 17%, “unsatisfied”: 8% Complication: $n = 5$ (20%) Re-operations: $n = 1$ (4.0%)</p>
Chung et al. [32]	<p><i>Indication/Diagnosis:</i> DDD, symptomatic</p> <p><i>Inclusion criteria</i> symptomatic DDD at L3–S1 confirmed by radiology, primary complaint of back pain, failed ≥ 6 months of conservative treatment, age between 18 and 60 years, minimum disc height of 4 mm, ODI ≥ 40</p> <p><i>Exclusion criteria</i> scoliosis, spondylolysis, spondylolisthesis, severe facet degeneration, osteoporosis</p>	<p>$N = 36$ ♀20 ♂16 mean 43 years (range 25–58 years) Type: ProDisc II® Level: 1-level: ($n = 25$) 2-level: ($n = 11$) L5/S1 ($n = 21$), L4/L5 ($n = 24$) L3/L4 ($n = 2$)</p>	<p>Follow up: 24 months (6 weeks, 12, 24 months), 5.3% lost to follow up ($n = 38$ originally included) Low back pain (VAS): pre-op: 75, follow up: 30, difference: –45 Leg pain (VAS): pre-op: 47, follow up: 12, difference: –35 Functional status (ODI): pre-op: 69.2%, follow up: 21.0%, difference: –48.2% Clinical success (ODI improvement > 25%): 94.4% ($n = 34$) Clinical success (ODI improvement > 75%): 44.4% ($n = 16$) Complication: $n = 5$ (13.8%) Re-operations: $n = 0$</p>
Siepe et al. [33, 34], Mayer et al. [36]	<p><i>Indication/Diagnosis:</i> DDD, symptomatic</p> <p><i>Inclusion criteria</i> Symptomatic DDD confirmed by radiology, primary complaint of low back pain, failed conservative treatment</p> <p><i>Exclusion criteria</i> DDD + disc herniation, post-discectomy, DDD + modic changes, central or lateral spinal stenosis, facet joint arthrosis/symptomatic facet joint problems, spondylolysis/spondylolisthesis, spinal instability (iatrogenic/altered posterior elements, e.g., following, laminectomy), major deformity/curvature deviations (e.g., scoliosis), metabolic bone disease (e.g., manifest osteoporosis/osteomalacia), previous operation with severe scarring and radiculopathy, compromised vertebral body (irregular endplate shape), previous/latent infection, metal allergy, spinal tumor, post-traumatic segments</p>	<p>$N = 99$ ♀60 ♂39 mean 43 years (range 22–66 years) Type: ProDisc II® Level: 1-level L5/S1 ($n = 57$), L4/L5 ($n = 22$), 2-level L4/L5/S1 ($n = 20$)</p>	<p>Follow up: 12 months (3 m, 6 m, 12 m), 0% lost to follow up ($n = 218$ included, 99 had reached 12 months FU) pain (VAS): pre-op: 70, follow up: 30, difference: –40 (SD 24) Functional status (ODI): pre-op: 42.0%, follow up: 21.0%, difference: –21.0% (SD 17) Return to work: 66.3% Patient Satisfaction: “completely satisfied” 58.5%, “satisfied” 20.2%, “unsatisfied” 21.3% Complication: $n = 17$ (17.2%) Re-operations: $n = 8$ (8.1%)</p>

Table 3 continued

Study/author	Indication/in/exclusion criteria	Demografy/type of disc for TDR/Level of intervention	Follow up: duration and outcome
Ogon et al. [35]	<i>Indication/Diagnosis:</i> degenerative disc disease (DDD) <i>Exclusion criteria</i> spondylarthrosis, spondylolisthesis, spinal stenosis, scoliosis, osteoporosis, infection, tumor	$N = 34 \text{ ♀} 26 \text{ ♂} 8$ mean 44 years (range 30–60 years) Type: ProDisc® Level: 1-level L5/S1 ($n = 22$), L4/L5 ($n = 10$), 2-level L4/L5/S1 ($n = 2$)	Follow up: 12 months (3 m, 12 m), lost to follow up not mentioned pain (VAS): pre-op: 74, follow up: 28, difference: −46 Functional status (ODI): pre-op: 48.0%, follow up: 13.1%, difference: −34.9% Functional status (SF-36) PCSS: pre-op: 31.3, follow up: 44.2, difference: −10.9 MCSS: pre-op: 38.6, follow up: 44.8, difference: −6.2 Patient Satisfaction: “completely satisfied”: 76.5%, “satisfied”: 14.7%, “unsatisfied”: 8.8% Complication: $n = 4$ (11.8%) Re-operations: $n = 1$ (2.9%)
Maverick® Huec et al. [37–39]	<i>Indication/Diagnosis:</i> DDD, symptomatic <i>Inclusion criteria:</i> symptomatic DDD confirmed by radiography and MRI, predominant chronic back pain and absence of permanent nerve root compression, failed ≥ 12 months of conservative treatment, ODI $> 30\%$, age between 20 and 60 years irrespective of sex <i>Exclusion criteria:</i> Severe osteoporosis, physiological dysfunction, history of previous infection, severe posterior element pathologies, fracture of the vertebra, tumor.	$N = 64 \text{ ♀} 39 \text{ ♂} 25$ mean 44 years (range 37–51 years) Intervention: TDR TDR + fusion Type: Maverick® Level: 1-level: L5/S1 ($n = 35$), L4/L5 ($n = 14$), L3/L4 ($n = 2$) 2-level: L5/S1 TDR + L5/S1 fusion ($n = 13$)	Follow up: 24 months (1 m, 3 m, 6 m, 12 m, 24 m), 0% lost to follow up Low back pain (VAS): pre-op: 76 (SD17), follow up: 32 (SD18), difference: −44, Leg pain (VAS): pre-op: 39, follow up: 21, difference: −18 Functional status (ODI): pre-op: 43.8% follow up: 23.1%, difference: −20.7% Clinical success (ODI improvement $> 25\%$): 75% Complication: $n = 26$ (40.6%) Reoperation: $n = 0$ (0%)
Acroflex® Fraser et al. [40]	<i>Indication/Diagnosis:</i> DDD, symptomatic <i>Inclusion criteria:</i> 1-level or 2-level symptomatic DDD at L4–S1 confirmed by radiology, back and/or leg (radicular) pain, failed ≥ 6 mo of conservative treatment, age between 30 and 55 years <i>Exclusion criteria:</i> previous lumbar surgery with the exception of discectomy and chemonucleolysis at the target level, steep lumbosacral angle at the target, central or lateral spinal stenosis, spondylolisthesis, systemic disease that would impact the patient's condition, obesity, alcohol and/or drug abuse, presence of three or more positive Waddell signs, scoliosis, major psychiatric disorder, current involvement in litigation related to the spinal condition or involvement in pursuing legal action related to the spinal condition	$N = 28 \text{ ♀} 14 \text{ ♂} 14$ mean 41 years (range 30–54 years) Type: Acroflex® Level: 1-level L5/S1 ($n = 19$), L4/L5 ($n = 5$), 2-level: L4/L5/S1 ($n = 4$)	Follow up: 24 months (6, 12 weeks, 6, 12, 24 months), lost to follow up not mentioned Low back pain (LBOS): pre-op: 17.7, follow up: 33.0, difference: −15.3 Functional status (ODI): pre-op: 49.3% follow up: 34.4%, difference: −14.9% Complication: $n = 17$ (60.7%) Re-operation: $n = 8$ (28.6%)

VAS (scale 0–100), ODI (scale 0–100)

All ODI scores and VAS scores were converted into 0–100 scale

primary outcome measures between the Charité group and the BAK cage at the 2-year follow up, and there is very low quality evidence (based on 1 study only with a high risk of bias) that there are no clinical relevant differences on the primary outcome measures at the 5-year follow up.

ProDisc® trial (Table 1) The ProDisc trial [52], which had a high risk of bias, randomized 236 patients to either TDR with the ProDisc device ($n = 161$) or to anterior lumbar circumferential fusion (using femoral ring allograft

and posterolateral fusion with autogenous iliac crest bone graft in combination with pedicle screws) ($n = 75$). Outcomes were reported with 2-year follow-up. Clinical success was defined using a combination of 10 outcomes as required by the FDA (Oswestry ≥ 15 points, SF-36 improvement, device success, neurologic success and six radiographic outcomes: no migration, no subsidence, no radiolucency, no loss of disc height, fusion status and ROM). Clinical success was statistically significantly better in the ProDisc (54.3%) than the fusion group (40.8%)

Table 4 Overview complication prospective cohort studies

	Maverick®			Acroflex®			ProDisc®			Charité®					Total range (%)	
	Huec et al. [37–39] (N = 64)	Fraser et al. [40] (N = 28) ♂14	Bertagnoli et al. [27] (N = 108)	Tropiano et al. [28] (N = 53)	Tropiano et al. [29] (N = 55, ♂30)	Bertagnoli et al. [30] (N = 104, ♂47)	Bertagnoli et al. [31] (N = 25, ♂15)	Chung et al. [32] (N = 36)	Siepe et al. [33] (N = 99, ♂39)	Ogon et al. [35] (N = 34)	Zeegers et al. [21] (N = 50, ♂20)	Regan [22] (N = 100)	Lemaire et al. [23] (N = 100, ♂41)	Ross et al. [24] (N = 160, ♂62)	Gioia et al. [25] (N = 36, ♂13)	Warachit [26] (N = 43)
Approach-related	11 (17.2%)	1 (7.1%) ^a			4 (8.7%) ^a	1 (2.1%) ^a	1 (6.7%) ^a	2 (5.5%)	3 (6.1%) ^a	2 (5.9%)	4 (11.0%) ^a	3 (4.5%) ^a	3 (4.5%) ^a	22 (18.7%) ^a	2 (10.5%) ^a	2.1–18.7
Prosthesis related		11 (39.3%)		3 (5.7%)			3 (12.0%)	2 (2.0%)			1 (2.0%)	5 (5.0%)	2 (2.0%)	9 (5.6%)	2 (5.6%)	2.0–39.3
Treatment related	14 (21.9%)	4 (14.3%)		2 (3.8%)	5 (9.0%)	4 (3.8%)		3 (8.3%)	11 (11.1%)	1 (2.9%)	31 (62.0%)	4 (4.0%)	10 (10.0%)	3 (1.9%)	3 (8.3%)	1.9–62.0
General surgical related	1 (1.6%)	1 (3.6%)	1 (1.0%)		1 (1.8%)		1 (4.0%)		1 (1.0%)	1 (2.9%)	7 (14.0%)		1 (1.0%)	19 (11.9%)	4 (11.1%)	1.0–14.0
Total of complications	26 (40.6%)	17 (64.3%)	1 (1.0%)	5 (9.4%)	10 (19.5%)	5 (5.9%)	5 (22.7%)	5 (13.8%)	17 (20.2%)	4 (11.8%)	44 (91.0%)	9 (9.0%)	16 (17.5%)	53 (38.1%)	11 (35.5%)	1.0–91.0
Reoperation at index level		8 (28.6%)		3 (5.7%)	3 (5.5%)	1 (1.0%)	1 (4.0%)		8 (8.1%)	1 (2.9%)	7 (14.0%)	7 (7.0%)	5 (5.0%)	12 (7.5%)	5 (13.9%)	1.0–28.6

^a Retrograde ejaculation calculated for men only

($P < 0.05$). Although this trial was designed as a non-inferiority study, it is unclear what statistical testing is applied. However, there were no significant differences between both groups on the mean functional impairment (−28.9 vs. −22.9%) and pain intensity scores (−39 vs. −32). In conclusion, there is very low quality evidence (based on 1 study only with a high risk of bias and inconsistent findings) for contradictory results on the primary outcome measures at the 2-year follow up for the ProDisc when compared with anterior lumbar circumferential fusion.

Flexicore® trial (Table 1) The Flexicore trial [53], with a high risk of bias reported the initial results of 76 patients from two clinics involved in a randomized multicentre controlled trial comparing the Flexicore device ($n = 44$) versus anterior lumbar circumferential fusion ($n = 23$) with 2 year follow-up. These 76 patients are only a small proportion of all randomized patients ($n = 401$) included in the complete trial. Overall, dropout rate was high, 33 patients (75%) in the intervention group and 16 patients (70%) in the control group after two years. Improvement in pain intensity (VAS −70 vs. −62) and functional impairment (Oswestry −56 vs. −46%) was slightly better in the Flexicore group than in the fusion group, but the authors did not report whether this difference was statistically significant or not. Because these are preliminary results, in addition to the high risk of bias, we refrain from drawing conclusions based on this study. In general, these results suggest no clinical relevant differences between TDR surgery and fusion techniques and a small overall success rate in both groups (approximately 50%).

(3) What is the safety of total disc replacement surgery?

Although some studies reported no major complications, other cohort studies describe a wide range (1.0–91.0%) of complication rates following TDR. The majority of these studies reported complication rates ranging from 10 to 40%. Complications can be separated into those related to the surgical approach (e.g. vascular injury, nerve root damage, retrograde ejaculation) range from 2.1 to 18.7%, related to the prosthesis (e.g. subsidence, migration, implant displacement, implant failure, end plate fracture) range from 2.0 to 39.3% and related to the treatment (e.g. wound, pain, neuromusculoskeletal) range from 1.9 to 62.0%. General surgical related complications ranged from 1.0 to 14.0%. Reoperation at index level was seen in 1.0–28.6% (Table 4). These reported complication rates and reoperation rates have to be interpreted carefully, because they have been described poorly.

Below we will describe the complications rates and reoperation rates as found in the three trials. The Flexicore

Table 5 Overview complications trials

	Prodisc®		Charité®		FlexiCore®	
	Index (<i>n</i> = 162), 82 ♂	Control (<i>n</i> = 80), 34 ♂	Index (<i>n</i> = 205), 113 ♂	Control (<i>n</i> = 99), 44 ♂	Index (<i>n</i> = 44)	Control (<i>n</i> = 23)
Approach-related	2 (2.4%) ^a	2 (2.5%)	16 (9.1%) ^a	9 (12.8%) ^a	1 (2.3%)	0
Prosthesis related	5 (3.1%)	0	8 (3.9%)	10 (10.1%)	2 (4.5%)	0
Treatment related	1 (0.6%)	0	17 (8.3%)	23 (23.2%)	5 (11.4%)	8 (34.8%)
General surgical related	2 (1.2%)	3 (3.8%)	16 (7.8%)	4 (4.0%)	2 (4.5%)	2 (8.7%)
Total of complications	10 (7.3%)	5 (6.3%)	57 (29.1%)	46 (50.2%)	10 (22.7%)	10 (43.5%)
Reoperation at index level	6 (3.7%)	4 (5.4%)	11 (5.4%)	9 (9.1%)	5 (11.4%)	6 (26.1%)
			13 (6.3%)	10 (10.1%)		

Table based on the published manuscripts, number of compilation in FDA report [56, 58] deviated (see text)

^a Retrograde ejaculation calculated for men only

trial [53] report 22.7% complications in the TDR group and 43.5% in the fusion group. Reoperations are reported in both groups; 11.4% for TDR and 26.1% for fusion (Table 5).

In the Charité trial, overall complication rates published by Blumenthal et al. [10] were 29.1% for TDR and 50.2% for fusion at 2-year follow-up. Device failures necessitating reoperation were reported in 5.4% of patients in de TDR group and 9.1% of patients in the fusion group at 2 year follow-up (Table 5). However, in the FDA report on the Charité trial much higher scores of adverse effects (TDR group 181.9% and fusion group 189.6%) were reported [56]. In an article from McAfee et al. [57], analysing the incidence of reoperations, even higher reoperation rates in the Charité trial are reported (6.3% in de TDR group and 10.1% in the fusion group).

In the ProDisc trial, there was a similar discrepancy between the article and the FDA report. The overall complication rate as reported by Zigler et al. [52] were 7.3% and 6.3% for TDR and fusion, respectively, but in the FDA report on the ProDisc trial much higher scores on adverse events were reported (TDR group 255.5% and fusion group 270.7%) [58]. Reoperation was necessary for 3.7% TDR patients and 5.4% fusion patients according to Zigler et al. The number of patients needed a reoperation was similar in the FDA report; however, the included number of patients in the trial was higher so the percentage of reoperation in the FDA rapport was slightly higher (Table 5).

Geisler et al. [59] analysed only the neurological complications in the Charité trial. The incidence was no higher in patients with the Charité (16.6%) than patients with BAK fusion (17.2%) ($P > 0.3$). Major neurologic complications in the Charité group (e.g. burning or dysesthetic leg pain, motor deficit in index level, nerve root injury) were reported in 4.9% and in the fusion group (e.g. burning or

dysesthetic leg pain, motor deficit at the index level) in 4%. One device related major complication, nerve root injury, was reported in the TDR group.

Leary et al. [60] reported on 18 patients requiring an anterior revision procedure for repositioning or removal of the Charité prosthesis because of complications. Three patients required revision of two levels. One patient had both levels revised in a single procedure, whereas two patients required staged procedures in order to revise both implants. Therefore, 21 implants were revised via 20 anterior procedures in 18 patients. Six revision cases were performed within the early postoperative period (7–14 days), all as a result of implant migration or dislocation. Late revision cases were required in 14 cases (range 3 weeks–4 years) due to implant migration, dislocation, end plate fractures, subsidence or persistent low back pain.

Van Ooij et al. [61–65] reported in several publications patients following implantation of the Charité prosthesis who experienced complications. Over the last 10 years, 75 patients with persisting back and leg pain and being unsatisfied with their clinical condition have been seen and analysed. An overview on late complications after TDR: subsidence ($n = 39$), prosthesis too small ($n = 24$), adjacent disc degeneration ($n = 36$), degenerative scoliosis ($n = 11$), facet joint degeneration on CT scan ($n = 25$), anterior migration ($n = 6$), posterior migration ($n = 2$), breakage metal wire ($n = 10$), wear ($n = 5$), severe osteolysis ($n = 1$), subluxation PE core ($n = 1$). 46 out of these 75 needed one or more salvage operations after their TDR. Fifteen patients were receiving posterior fusion without removal of the prosthesis. Because of persisting pain, afterwards 4 patients had their prosthesis removed in an additional operation. In 22 patients, 26 prostheses were removed and an anterior and posterior fusion was performed. In addition, seven patients received posterior

fusion elsewhere, and in two patients, the disc prosthesis was removed elsewhere. Intraoperatively, the surgeon encountered three times vessel damage. In conclusion, a wide range of complications rates following TDR (1–90.0%) was found in all cohort studies. The majority of the studies reported complication rates ranging from 10 to 40%. Reoperation at index level was reported in 1.0–28.6%. The three randomized controlled trials published overall complication rates range from 7.2 to 28.6% in the TDR group and 6.7 to 50.2% in the fusion group. The overall reoperation rate at the index level ranged from 3.7 to 11.4% in the TDR group and 5.4–26.1% in the fusion group. However, much higher rates were reported in FDA reports on the Charité and ProDisc trials.

Discussion

In this article, we systematically reviewed the available literature on the clinical course, effectiveness, cost-effectiveness, and safety of TDR in patients with symptomatic DDD. Sixteen prospective cohort studies were identified that assessed the course of complaints and symptoms. These studies suggest pain relief, improvement in functional status and patient satisfaction after TDR. However, the quality of reporting on outcomes was often poor, hampering an adequate interpretation. In addition, a significant amount of complications was reported. These cohort studies lacked control group, which is necessary to evaluate effectiveness of TDR. Only three randomized controlled multicentre trials were identified that had assessed the effectiveness of TDR. The results show that there is low quality evidence (based on one study only with a low risk of bias) that there are no clinically relevant differences on the primary outcome measures between the Charité group and the BAK cage at 2 years follow up, and there is very low quality evidence (based on 1 study only with a high risk of bias) that there are no clinically relevant differences on the primary outcome measures at 5 years follow up. Furthermore, there is very low quality evidence (based on one study only with a high risk of bias) for contradictory results on the primary outcome measures for the ProDisc when compared with anterior lumbar circumferential fusion at the 2-year follow up. There is insufficient evidence on the Flexicore, because this trial had a high risk of bias, and should be considered as a preliminary report because it only reported on a small proportion of all included patients who participated in this multi centre trial.

For assessing the complication rate, all reported complications were extracted from the cohort studies and randomized controlled trials included in this review, as well as overview studies on complication rates. A wide range of complications rates following TDR (1–91.0%) was found

in the cohort studies. The majority of the studies reported complication rates ranging from 10 to 40%. Reoperation at the index level was reported in 1.0–28.6%. In the three randomized controlled trials, overall complication rates ranged from 7.3 to 29.1% in the TDR group and from 6.3 to 50.2% in the fusion group. The overall reoperation rate at the index level ranged from 3.7 to 11.4% in the TDR group and from 5.4 to 26.1% in the fusion group. However, much higher rates were reported in FDA reports on the Charité and ProDisc trials. No full economic evaluation was identified, so there is no evidence regarding the cost-effectiveness of TDR.

The course of DDD complaints and/or symptoms following total disc replacement surgery

We identified 16 prospective cohort studies to evaluate the course of DDD complaints and/or symptoms. The outcome results suggested a positive course after TDR with a high proportion of patients satisfied with the result. However, these studies were of poor methodological quality and detailed information on how outcomes were measured was often lacking. For example, it was often unclear which criteria were used for clinical success and how return to work was measured. Furthermore, another drawback is that a significant amount of complications was reported as well.

Moreover, these results have to be interpreted in light of controversy and limited literature regarding the causal relationship between DDD and chronic low back pain [4]. Boden et al. [66] reported on 67 asymptomatic individuals assessed for DDD with MRI. DDD was seen in 34% of the individuals between 20 and 39 years of age; 59% of individuals between 40 and 59 years of age, and in all but one (93%) between 60 and 80 years of age. Jensen et al. [67] reported on 98 asymptomatic people after MRI and concluded that 64% of these people had an intervertebral disc abnormality. This challenges the rationale of surgery for DDD in the absence of convincing pathological pathways of DDD.

The effectiveness of total disc replacement surgery compared to other treatments

The Flexicore trial should be interpreted with great caution because of the high risk of bias. Of the three randomized controlled trials, the 2-year follow up of the Charité trial was considered to have a low risk of bias. However, the fusion technique with BAK cages and the iliac crest bone autograft used in this trial are techniques that are no longer commonly used because of poor outcomes [68–70]. A better comparator would be the circumferential fusion technique which was used in the ProDisc and Flexicore trials. The use of autograft in all three studies may also be

criticized as many surgeons now use both recombinant BMP-2 and/or percutaneous pedicle screw fixation when performing lumbar fusion [55]. Use of an inadequate control intervention brings into question the clinical relevance of the results of the three trials. An additional concern is the fact that the literature is still controversial about the superiority of fusion compared to conservative treatment [5, 15, 71, 72]. For this reason, it can be interesting to compare the effectiveness of TDR to conservative treatment. At present, no studies comparing total disc replacement surgery to other treatments have been published.

The three randomized controlled trials selected patients carefully, scrutinizing various contraindications for TDR. Because of this careful selection, the published trials do not provide evidence for the widespread use of TDR in all patients with DDD. The relevance of the clinical outcomes in the Charité and the ProDisc trials can also be challenged. First, modest success rates were observed in both the TDR and the fusion groups. In the Charité trial, only 57.1% of patients with TDR met all 4 criteria for success, when compared with 46.5% in the fusion group ($P < 0.0001$). In the ProDisc group, only 53.4% of patients with TDR met all 10 FDA criteria for success, when compared with 40.8% in the fusion group ($P = 0.0438$). Second, in the Prodisc trial, 69.1% of TDR subjects improved by more than 25% on the Oswestry, when compared with 54.9% in the fusion group. In the Charité trial, 63.9% of TDR subjects improved by more than 25% on the Oswestry, compared to 50.5% in the fusion group. The use of the 25% benchmark for improvement should be interpreted against a background of a recently published consensus statement that advocates a 30% improvement in Oswestry as a benchmark for clinically relevant improvement. This recommendation focussed primarily on conservative interventions in a primary care setting. It was suggested that it might be more appropriate to use larger change scores as benchmarks for expensive and risky procedures [73]. Third, one of the purposes of the device implementation is to reduce low back pain whereas the definition of success did not consider pain relief or opioid use. Finally, Oswestry and VAS cannot discriminate between pain that is residual from the iliac crest after fusion surgery versus the lumbar spine. Therefore, Oswestry and VAS may be artificially higher in the fusion group compared with TDR.

The ODI was used in all included RCT's, but different versions of the ODI were used. Sasso used ODIv2.0 [53]. Blumenthal used the ODIv1.0 and Zigler used the ODI (chiropractic revised version [74]) [75]. Because different versions of the ODI are used, a direct comparison between studies is hampered. Zigler, however, holds the opinion that the differences between the various ODI versions are subtle and, they think, inconsequential [77].

Davidson [77] and Fairbanks [75] hold the opinion that the amendments of this 'chiropractic revised version' are major and therefore this version cannot be compared with the official versions of ODI.

The safety of the total disc replacement surgery

Complications have been poorly described in the prospective cohort studies and the randomized controlled trials. It is interesting that the complications rates and reoperation rates are lower in the published articles than in the FDA reports [56, 58]. This illustrates the complexity of reporting on adverse effects. Compared to the journals where the papers were published, apparently the FDA requires exhaustive and detailed reporting of "adverse events" most of which have no relationship to the success or failure of the prosthesis. Complications associated with lumbar fusion include incomplete relief of pain, loss of motion, loss of sagittal balance, pseudoarthrosis, adjacent segment degeneration, and bonegraft donor site complication. However, a separate set of concerns exist in TDR. Wear debris leading to osteolysis and systematic effects, vertebral body damage, posterior migration or extrusion may lead to device failure and serious vascular complications. Prosthesis that fail to adequately replicate the physiologic kinematics of the lumbar spine may predispose the patient to facet joint degeneration. Without true motion preservation, the devices will merely act as interbody spacers with no potential to prevent adjacent level degeneration [78]. Finally, reported complications for TDR show there can be severe and even life threatening, e.g. major vascular injury, major nerve root damage and device failure. However, these complication rates are low [9, 17, 79, 80].

Furthermore, in the two low risks of bias studies [10, 52], the re-operation rates in the TDR group are slightly higher than in the fusion group. However, this has to be balanced against the fact that re-operation procedures for TDR are more complex.

The use of intervertebral disc prostheses as an alternative to spinal fusion has been advocated to preserve segmental motion and to prevent adjacent degeneration. However, there is no consensus on this subject in literature. Some studies suggest adjacent level degeneration is prevented after TDR [6, 12]. However, other studies show adjacent disc degeneration after TDR [61, 81]. This could be the result of the DDD itself, spreading to multiple levels of the spine, and/or be a consequence of stresses on adjacent levels, generated from unphysiological motion and functioning of the disc prosthesis [61]. Moreover, there is little knowledge regarding complications on the long term. Putzier et al. [81] published a retrospective study with 17 years follow-up and reoperation was necessary in 11%

of patients. It is important to know more about long-term complications because most operated patients are of relatively young age, between 30 and 50 years. A disc prosthesis used for TDR should survive for at least 40 years. It is very questionable if the lifetime of the designs now available will be that long as little is known about long-term behaviour of biomaterials in the spine. We do know that revision surgery can be dangerous because of adherence to great vessels and the nerve plexus. Studies that review long-term complications and longevity of the prostheses are highly recommended.

Conclusion

There is low quality evidence that there are no clinically relevant differences on the primary outcome measures between the Charité group and the BAK cage at 2 years follow up, and there is very low quality evidence that there are no clinical relevant differences on the primary outcome measures at 5 years follow up. For the ProDisc device, there is very low quality evidence for contradictory results on the primary outcome measures when compared with anterior lumbar circumferential fusion. Furthermore, reported complication rates varied from 1.0 to 91.0% in cohort studies and 7.3 to 29.1% in randomized controlled trials. Still lacking are high quality prospective, controlled, long-term follow-up studies, including a full economic evaluation taking into account all relevant cost when compared with the clinical benefit, and with relevant control groups to establish the efficiency and the longevity of the devices. The existing evidence, specifically regarding long-term effectiveness and/or safety is considered insufficient to justify the widespread use of TDR over fusion for single level degenerative disc. It is recommended that disc replacement surgery at this time only is performed within prospective scientific studies until further documentation of its efficiency is provided.

Acknowledgments For this review, the authors received a grant from The Health Care Insurance Board (CVZ), Diemen, The Netherlands.

Conflict of interest statement For this review, the authors received a Grant from The Health Care Insurance Board (CVZ), Diemen, The Netherlands.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Van Tulder MW, Koes BW, Bouter LM (1995) A cost-of-illness study of back pain in The Netherlands. *Pain* 1995:233–240
2. Pengel LHM, Herbert RD, Maher CG et al (2003) Acute low back pain: systematic review of its prognosis. *BMJ* 327:323
3. Errico TJ (2005) Lumbar disc arthroplasty. *Clin Orthop Relat Res* 435:106–117
4. Frelinghuysen P, Huang RC, Girardi FP et al (2005) Lumbar total disc replacement part I: rationale, biomechanics, and implant types. *Orthop Clin North Am* 36:293–299
5. Fritzell P, Hagg O, Wessberg P et al (2001) Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 26:2521–2532
6. Hochschuler SH, Ohnmeiss DD, Guyer RD et al (2002) Artificial disc: preliminary results of a prospective study in the United States. *Eur Spine J* 2:S106–S110
7. Errico TJ (2004) Why a mechanical disc? *Spine J* 4:151S–157S
8. Freeman BJC, Davenport J (2006) Total disc replacement in the lumbar spine: a systematic review of the literature. *Eur Spine J* 15(Suppl 3):S439–S447
9. de Kleuver M, Oner FC, Jacobs WC (2003) Total disc replacement for chronic low back pain: background and a systematic review of the literature. *Eur Spine J* 12:S108–S116
10. Blumenthal S, McAfee PC, Guyer RD et al (2005) A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine* 30:1565–1575
11. Cunningham BW, Dmitriev AE, Hu N et al (2003) General principles of total disc replacement arthroplasty: seventeen cases in a nonhuman primate model. *Spine* 28:S118–S124
12. Link HD (2002) History, design and biomechanics of the LINK SB Charite artificial disc. *Eur Spine J* 11:S98–S105
13. McAfee PC, Cunningham B, Holsapple G et al (2005) A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. *Spine* 30:1576–1583
14. Szpalski M, Gunzburg R, Mayer M (2002) Spine arthroplasty: a historical review. *Eur Spine J* 11(Suppl 2):S65–S84
15. Fairbank J, Frost H, Wilson-MacDonald J et al (2005) Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *BMJ* 330:1233
16. Fernstrom U (1966) Arthroplasty with intercorpal endoprosthesis in herniated disc and in painful disc. *Acta Chir Scand Suppl* 357:154–159
17. Randolph GB, Scioscia TN, Wang JC (2006) Lumbar total disc arthroplasty: state of the data. *Semin Spine Surg* 18:61–71
18. van Tulder M, Furlan A, Bombardier C et al (2003) Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine* 28:1290–1299
19. Brozek J, Akl E, Alonso-Coello P, on behalf of the GRADE Working Group et al (2009) Grading quality of evidence and strength of recommendations in clinical practice guidelines. *Allergy* 64:669–677
20. Lemaire JP, Skalli W, Lavaste F et al (1997) Intervertebral disc prosthesis. Results and prospects for the year 2000. *Clin Orthop Relat Res* 337:64–76
21. Zeegers WS, Bohnen LM, Laaper M et al (1999) Artificial disc replacement with the modular type SB Charite III: 2-year results in 50 prospectively studied patients. *Eur Spine J* 8:210–217
22. Regan JJ (2005) Clinical results of charite lumbar total disc replacement. *Orthop Clin North Am* 36:323–340

23. Lemaire JP, Carrier H, Eh SariAli et al (2005) Clinical and radiological outcomes with the Charite artificial disc: a 10-year minimum follow-up. *J Spinal Disord Tech* 18:353–359
24. Ross R, Mirza AH, Norris HE et al (2007) Survival and clinical outcome of SB Charite III disc replacement for back pain. *J Bone Joint Surg Br* 89:785–789
25. Gioia G, Mandelli D, Randelli F (2007) The Charite III artificial disc lumbar disc prosthesis: assessment of medium-term results. *J Orthop Traumatol* 8:134–139
26. Warachit P (2008) Results of Charite artificial lumbar disc replacement: experience in 43 Thais. *J Med Assoc Thai* 91:1212–1217
27. Bertagnoli R, Kumar S (2002) Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. *Eur Spine J* 11(Suppl 2):S131–S136
28. Tropiano P, Huang RC, Girardi FP et al (2003) Lumbar disc replacement: preliminary results with ProDisc II after a minimum follow-up period of 1 year. *J Spinal Disord Tech* 16:362–368
29. Tropiano P, Huang RC, Girardi FP et al (2005) Lumbar total disc replacement. Seven to eleven-year follow-up. *J Bone Joint Surg Am* 87:490–496
30. Bertagnoli R, Yue JJ, Shah RV et al (2005) The treatment of disabling single-level lumbar discogenic low back pain with total disc arthroplasty utilizing the Prodisc prosthesis: a prospective study with 2-year minimum follow-up. *Spine* 30:2230–2236
31. Bertagnoli R, Yue JJ, Shah RV et al (2005) The treatment of disabling multilevel lumbar discogenic low back pain with total disc arthroplasty utilizing the ProDisc prosthesis: a prospective study with 2-year minimum follow-up. *Spine* 30:2192–2199
32. Chung SS, Lee CS, Kang CS (2006) Lumbar total disc replacement using ProDisc II: a prospective study with a 2-year minimum follow-up. *J Spinal Disord Tech* 19:411–415
33. Siepe CJ, Mayer HM, Wiechert K et al (2006) Clinical results of total lumbar disc replacement with ProDisc II: three-year results for different indications. *Spine* 31:1923–1932
34. Siepe CJ, Mayer HM, Heinz-Leisenheimer M et al (2007) Total lumbar disc replacement: different results for different levels. *Spine* 32:782–790
35. Ogon M, Howanietz N, Tuschel A et al (2007) Implantation of the Prodisc® intervertebral disk prosthesis for the lumbar spine. *Oper Orthop Traumatol* 19:209–230
36. Mayer HM, Wiechert K, Korge A et al (2002) Minimally invasive total disc replacement: surgical technique and preliminary clinical results. *Eur Spine J* 11(Suppl 2):S124–S130
37. Le Huec JC, Mathews H, Basso Y et al (2005) Clinical results of Maverick lumbar total disc replacement: two-year prospective follow-up. *Orthop Clin North Am* 36:315–322
38. Le Huec JC, Basso Y, Aunoble S et al (2005) Influence of facet and posterior muscle degeneration on clinical results of lumbar total disc replacement: two-year follow-up. *J Spinal Disord Tech* 18:219–223
39. Le Huec J, Basso Y, Mathews H et al (2005) The effect of single-level, total disc arthroplasty on sagittal balance parameters: a prospective study. *Eur Spine J* 14:480–486
40. Fraser RD, Ross ER, Lowery GL et al (2004) AcroFlex design and results. *Spine J* 4:245S–251S
41. McAfee PC, Fedder IL, Saiedy S et al (2003) Experimental design of total disk replacement experience with a prospective randomized study of the SB Charite. *Spine* 28:S153–S162
42. McAfee PC, Fedder IL, Saiedy S et al (2003) SB Charite disc replacement: report of 60 prospective randomized cases in a US center. *J Spinal Disord Tech* 16:424–433
43. Guyer RD, McAfee PC, Hochschuler SH et al (2004) Prospective randomized study of the Charite artificial disc: data from two investigational centers. *Spine J* 4:252S–259S
44. Blumenthal SL, Ohnmeiss DD, Guyer RD et al (2003) Prospective study evaluating total disc replacement: preliminary results. *J Spinal Disord Tech* 16:450–454
45. Zigler JE (2004) Lumbar spine arthroplasty using the ProDisc II. *Spine J* 4:260S–267S
46. Delamarter RB, Fribourg DM, Kanim LEA et al (2003) ProDisc artificial total lumbar disc replacement: introduction and early results from the United States clinical trial. *Spine* 28:S167–S175
47. Zigler JE (2003) Clinical results with ProDisc: European experience and U.S. investigation device exemption study. *Spine* 28:S163–S166
48. Zigler JE, Burd TA, Vialle EN et al (2003) Lumbar spine arthroplasty: early results using the ProDisc II: a prospective randomized trial of arthroplasty versus fusion. *J Spinal Disord Tech* 16:352–361
49. Delamarter RB, Bae HW, Pradhan BB (2005) Clinical results of ProDisc-II lumbar total disc replacement: report from the United States clinical trial. *Orthop Clin North Am* 36:301–313
50. Auerbach JD, Wills BPD, McIntosh TC et al (2005) Lumbar disc arthroplasty versus fusion for single-level degenerative disc disease: two-year results from a randomized prospective study. *Semin Spine Surg* 17:310–318
51. Guyer RD, McAfee PC, Banco RJ et al (2009) Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year followup. *Spine J* 9:374–386
52. Zigler J, Delamarter R, Spivak JM et al (2007) Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine* 32:1155–1162
53. Sasso RC, Foulk DM, Hahn M (2008) Prospective, randomized trial of metal-on-metal artificial lumbar disc replacement: initial results for treatment of discogenic pain. *Spine* 33:123–131
54. Mayer HM, Siepe C (2007) Total lumbar disc arthroplasty. *Curr Orthop* 21:17–24
55. German JW, Foley KT (2005) Disc arthroplasty in the management of the painful lumbar motion segment. *Spine* 30:S60–S67
56. FDA approval Charité: <http://www.fda.gov/cdrh/pdf4/P040006a.pdf>
57. McAfee PC, Geisler FH, Saiedy SS et al (2006) Revisability of the CHARITE artificial disc replacement: analysis of 688 patients enrolled in the U.S. IDE study of the CHARITE Artificial Disc. *Spine* 31:1217–1226
58. FDA approval ProDisc: <http://www.fda.gov/cdrh/pdf5/p050010a.pdf>
59. Geisler FH, Blumenthal SL, Guyer RD et al (2004) Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: results of a multicenter, prospective, randomized investigational device exemption study of Charite intervertebral disc. *J Neurosurg Spine* 1:143–154
60. Leary SP, Regan JJ, Lanman TH et al (2007) Revision and explantation strategies involving the CHARITE lumbar artificial disc replacement. *Spine* 32:1001–1011
61. Punt IM, Visser VM, van Rhijn LW et al (2008) Complications and reoperations of the SB Charite lumbar disc prosthesis: experience in 75 patients. *Eur Spine J* 17:36–43
62. van Ooij A, Oner FC, Verbout AJ (2003) Complications of artificial disc replacement: a report of 27 patients with the SB Charite disc. *J Spinal Disord Tech* 16:369–383
63. van Ooij A, Schurink GW, Oner FC et al (2007) Findings in 67 patients with recurrent or persistent symptoms after implantation

- of a disc prosthesis for low back pain. *Ned Tijdschr Geneesk* 151:1577–1584
64. Kurtz SM, van Ooij A, Ross R et al (2007) Polyethylene wear and rim fracture in total disc arthroplasty. *Spine J* 7:12–21
 65. van Ooij A, Kurtz SM, Stessels F et al (2007) Polyethylene wear debris and long-term clinical failure of the Charite disc prosthesis: a study of 4 patients. *Spine* 32:223–229
 66. Boden SD, Davis DO, Dina TS et al (1990) Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 72:403–408
 67. Jensen MC, Brant-Zawadzki MN, Obuchowski N et al (1994) Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 331:69–73
 68. Button G, Gupta M, Barrett C et al (2005) Three- to six-year follow-up of stand-alone BAK cages implanted by a single surgeon. *Spine J* 5:155–160
 69. Lorenz M, Zindrick M, Schwaegler P et al (1991) A comparison of single-level fusions with and without hardware. *Spine* 16:S455–S458
 70. Pellise F, Puig O, Rivas A et al (2002) Low fusion rate after L5–S1 laparoscopic anterior lumbar interbody fusion using twin stand-alone carbon fiber cages. *Spine* 27:1665–1669
 71. Brox JJ, Reikeras O, Nygaard O et al (2006) Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. *Pain* 122:145–155
 72. Van Tulder MW, Koes B, Seitsalo S et al (2006) Outcome of invasive treatment modalities on back pain and sciatica: an evidence-based review. *Eur Spine J* 15:S82–S92
 73. Ostelo RWJG, Deyo RA, Stratford P et al (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 33:90–94
 74. Hudson-Cook N, Tomes-Nicholson K, Breen A (1989) A revised Oswestry disability questionnaire. In: Roland M, Jenner JR (eds) *Back pain: new approaches to rehabilitation and education*. Manchester University Press, Manchester, pp 187–204
 75. Fairbank J (2007) Letter to the Editor. *Spine* 32(25):2929–2930
 76. Stieber JR, Donald GD (2006) Early failure of lumbar disc replacement: case report and review of the literature. *J Spinal Disord Tech* 19:55–60
 77. Zigler JE (2007) Letter to the editor. *Spine* 32(25):2930–2931
 78. Brau SA (2002) Mini-open approach to the spine for anterior lumbar interbody fusion: description of the procedure, results and complications. *Spine J* 2:216–223
 79. Bertagnoli R, Zigler J, Karg A et al (2005) Complications and strategies for revision surgery in total disc replacement. *Orthop Clin North Am* 36:389–395
 80. Harrop JS, Youssef JA, Maltenfort M et al (2008) Lumbar adjacent segment degeneration and disease after arthrodesis and total disc arthroplasty. *Spine* 33:1701–1707
 81. Putzier M, Funk JF, Schneider SV et al (2006) Charite total disc replacement—clinical and radiographical results after an average follow-up of 17 years. *Eur Spine J* 15:183–195